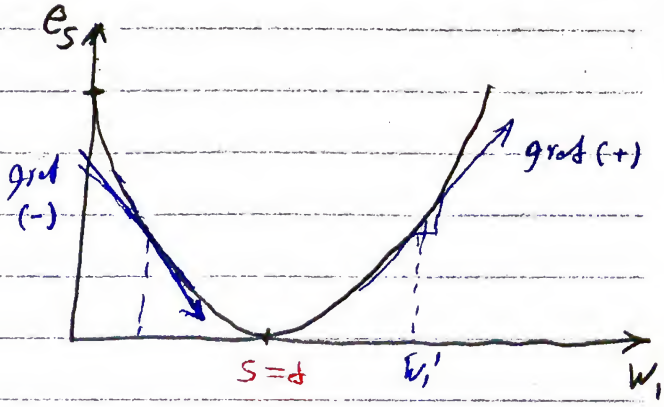
Back Propagation

((Gradient Desent))

$e = d - s$

$e_s = \frac{1}{2} e^2 = \frac{1}{2} (d-s)^2$

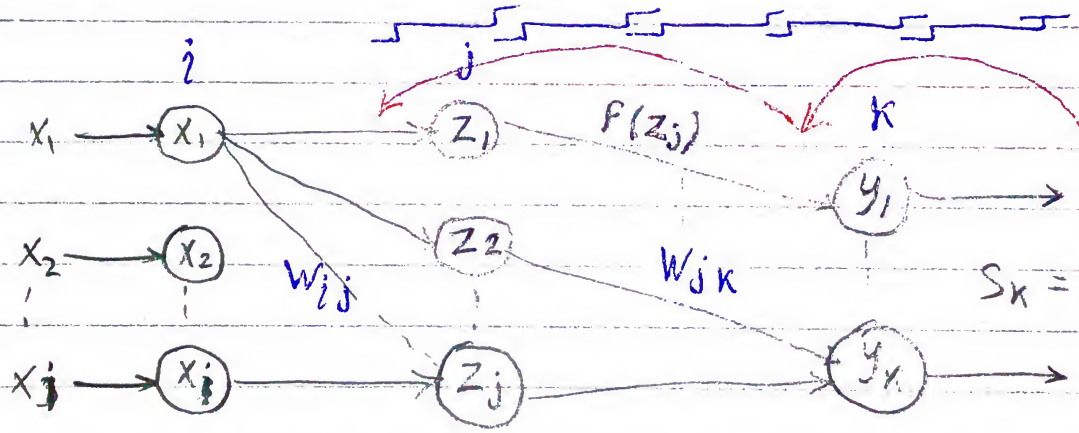


$\frac{\partial e_s}{w_1}$ = gradient

Back propagation يشترط ان يكون الـ Activation fn. قابلة للاشتقاق

$\Delta w \rightarrow$ positive | $\Delta w \rightarrow$ negative

Δw = grad



$\rightarrow F \rightarrow S = \checkmark$

$\rightarrow e = d - S = \checkmark$

1. Find the output for specific input
2. calculate the error
3. error back to previous layer to change weights

$S_k = F(y_k)$ | $z_j = \sum_i w_{ij} x_i$

$y_k = \sum_j w_{jk} F(z_j)$ | $E = \frac{1}{2} \sum_k (d_k - F(y_k))^2$

Update Weights :

$$\frac{\partial E}{\partial w_{jk}} = -\frac{2}{2}\left(d_k - f(y_k)\right)\frac{\partial f(y_k)}{\partial w_{jk}} \quad \text{(For output layer)}$$

$$= -e\,\frac{\partial f(y_k)}{\partial w_{jk}}$$

$$= -e\,\underbrace{\frac{\partial f(y_k)}{\partial y_k}}\cdot\frac{\partial y_k}{\partial w_{jk}}$$

depends on Activation Function

$$= -e\,\frac{\partial f(y_k)}{\partial y_k}\,f(z_j) \qquad \delta_k = e\,\frac{\partial f(y_k)}{\partial y_k}$$

For sigmoidal function

$$= -e\,f(y_k)\,(1-f(y_k))\,f(z_j)$$

$$\delta_k = e\,f(y_k)\,(1-f(y_k))$$

$$\frac{\partial E}{\partial w_{jk}} = -\delta_k\,f(z_j)$$

For weight from input to hidden layer

$$\frac{\partial E}{\partial w_{ij}} = \sum_k -\frac{2}{2}\left(d_k - f(y_k)\right)\frac{\partial f(y_k)}{\partial w_{ij}}$$

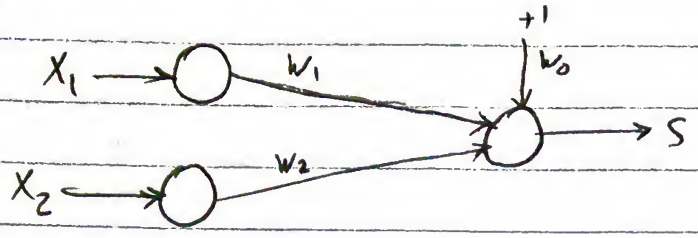
$$= \sum_k -e\,\underbrace{\frac{\partial f(y_k)}{\partial y_k}}\cdot\underbrace{\frac{\partial y_k}{\partial f(z_j)}}\cdot\underbrace{\frac{\partial f(z_j)}{\partial z_j}}\cdot\underbrace{\frac{\partial z_j}{\partial w_{ij}}}$$

$$= \sum_k -\delta_k \quad w_{jk} \quad \underbrace{f(z_j)\,(1-f(z_j))}_{\text{for sigmoid}} \quad x_i$$

$$= -\Big(\sum_k \delta_k w_{jk}\Big)\,f(z_j)\,(1-f(z_j))\,x_i$$

$$= -\delta_j\,x_i$$

(1)



$w_1 = -0.5$ $\quad w_2 = 0.5$ $\quad w_0 = 0.6$

$d = 0.9$ $\quad X_1 = 3.1$ $\quad X_2 = 2.7$ $\quad$ Sigmoidal

a) Find e, E

b) Find gradient for w_0, w_1, w_2

a) $S = f(y) = ?$

$y = X_1 w_1 + X_2 w_2 + w_0 = 0.4$

$S = f(y) = \frac{1}{1+e^{-y}} = 0.599$

$e = d - S = 0.9 - 0.599 = 0.301$

$E = \frac{1}{2}(d-S)^2 = \frac{1}{2} e^2 = 0.045$

b) $\frac{\partial E}{\partial w_i} = \frac{\partial E}{\partial f(y)} \cdot \frac{\partial f(y)}{\partial y} \cdot \frac{\partial y}{\partial w_i}$

$\frac{\partial E}{\partial f(y)} \cdot \frac{\partial f(y)}{\partial y} = -\delta_K$ $\quad \frac{\partial y}{\partial w_i} = X_i$

$\delta_K = e f(y)(1 - f(y)) = 0.072$

$\frac{\partial E}{\partial w_0} = -\delta_K = -0.072$

$\frac{\partial E}{\partial w_1} = -\delta_K X_1 = +0.229$

$\frac{\partial E}{\partial w_1} = -\delta_K X_2 = -0.194$

[2] For Bipolar Sigmoid, from (1)

$y = 0.4$

$g(y) = \frac{2}{1+e^{y}} - 1 = 0.197$

$\frac{\partial E}{\partial w_i} = \frac{\partial E}{\partial g(y)} \cdot \frac{\partial g(y)}{\partial y} \cdot \frac{\partial y}{\partial w_i}$

$\delta_k = \epsilon \cdot \frac{1}{2}(1 - g(y)^2) = 0.338$

$\frac{\partial E}{\partial w_0} = -\delta_k = -0.338$

$\frac{\partial E}{\partial w_1} = -\delta_k x_1 = -1.048$

$\frac{\partial E}{\partial w_2} = -\delta_k x_2 = -0.913$

[3] from (1)

$\frac{\partial E}{\partial w_0} = -0.13 \qquad \frac{\partial E}{\partial w_1} = -0.52 \qquad \frac{\partial E}{\partial w_2} = -0.65$

Find x_1, x_2 $\quad w_0 = 0.6 \quad w_1 = -0.5 \quad w_2 = 0.5$

$\delta = 0.9$

$\frac{\partial E}{\partial w_0} = -\delta_k = -0.13$

$\delta_k = 0.13$

$\frac{\partial E}{\partial w_1} = -\delta_k x_1$

$-0.13 x_1 = -0.52 \rightarrow x_1 = 4$

$\frac{\partial E}{\partial w_2} = -\delta_k x_2 \Rightarrow x_2 = 5$

4 from (3)

$w_1 = 0.5$ $\quad w_2 = -0.5$ $\quad e = 0.52$

find w_0

$\delta_k = 0.13$

Sigmidal

~~e - s~~

~~s - e~~

$$\delta = e \; F(y)\,(1 - F(y))$$

$$0.13 = 0.52\,(S - S^2)$$

$$\frac{0.13}{0.52} = S - S^2$$

$$S^2 - S + 0.25 = 0$$

$$(S - 0.5)^2 = 0$$

$$S = F(y) = 0.5$$

$$y = \ln \frac{F(y)}{1 - F(y)} = 0$$

$$y = w_1 x_1 + w_2 x_2 + w_0$$

$$0 = 0.5 * 4 + 0.5 * 5 + w_0$$

$$\boxed{w_0 = 0.5}$$

6 to part (4) → $e = 0.5$

$$F(y)\,(1 - F(y)) = \frac{\delta}{e} = \frac{0.13}{0.5} = 0.26$$

$$F(y) = \frac{1 \pm \sqrt{1 - 4(0.26)}}{2}$$

no Solution

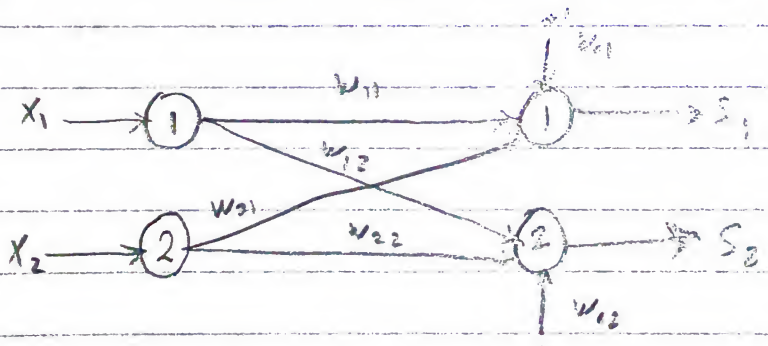
$$\frac{\delta}{e} \leq 0.25$$

$$e \geq \frac{0.13}{0.25}$$

$$e \geq 0.52$$

أقل Error يمكن الوصول إليه

7



$X_1 = -2.6$ $w_{01} = 0.35$ $w_{11} = 1.1$ $w_{21} = -0.8$

$X_1 = -1.9$ $w_{02} = -0.46$ $w_{12} = -1.2$ $w_{22} = 0.7$

$d_1 = 0.51$ $d_2 = 1.21$

Req a) e_1, e_2

b) gradient of all weights

a) $e_1 = d_1 - s_1 = 0.51 - s_1$

$s_1 = f(y_1)$

$y_1 = -0.99$

$s_1 = \frac{1}{1+e^{-y}} = 0.271$

$e_1 = 0.51 - 0.271 = 0.239$

$y_2 = 1.33$

$s_2 = 0.791$

$e_2 = 0.419$

$E = \frac{1}{2}\left[(d_1 - s_1)^2 + (d_2 - s_2)^2\right]$

$= \frac{1}{2}(e_1^2 + e_2^2)$

$= 0.116$

b) Connections to output neuron (1)

(w_{01}, w_{11}, w_{21})

$\frac{\partial E}{\partial w_{01}} = \frac{\partial E}{\partial f(y_1)} \cdot \frac{\partial f(y_1)}{\partial y_1} \cdot \frac{\partial y_1}{w_{01}} = -\delta_1 x_0$

$\delta_1 = e_1 s_1 (1 - s_1) = 0.047$

$\frac{\partial E}{\partial w_{01}} = -\delta_1 = -0.047$

$\frac{\partial E}{\partial w_{11}} = -\delta_1 x_1 = 0.122$

$\frac{\partial E}{\partial w_{21}} = -\delta_1 x_2 = 0.089$

Connections to output neuron ②
(w_{02}, w_{12}, w_{22})

$$\delta_2 = e_2 S_2(1 - S_2) = 0.069$$

$$\frac{\partial E}{\partial w_{02}} = -\delta_2 = -0.069$$

$$\frac{\partial E}{\partial w_{12}} = -\delta_2 x_1 = 0.179$$

$$\frac{\partial E}{\partial w_{22}} = -\delta_2 x_2 = 0.131$$

[8] Pro ① Bipolar Sigmoid

$e_1 = 1.22$ $\quad d_1 = 1.41$

$e_2 = -0.81$ $\quad d_2 = -1.23$

$w_{01} = -0.53$ $\quad w_{11} = -0.92$ $\quad w_{21} = 0.76$

$w_{02} = 0.53$ $\quad w_{12} = 0.87$ $\quad w_{22} = -0.65$

a) Find x_1, x_2

b) Find gradient

a) $e_1 = d_1 - s_1 \longrightarrow s_1 = 0.19$

$e_2 = d_2 - s_2 \longrightarrow s_2 = -0.42$

$$y_1 = \ln \frac{1+s_1}{1-s_1} = 0.385$$

$$y_2 = \ln \frac{1+s_2}{1-s_2} = -0.895$$

$y_1 = -0.92 x_1 + 0.76 x_2 - 0.53 = 0.385$

$0.92 x_1 - 0.76 x_2 = -0.915 \rightarrow ①$

$y_2 = 0.87 x_1 - 0.65 x_2 + 0.53 = -0.895$

$0.87 x_1 - 0.65 x_2 = -1.425 \rightarrow ②$

* Solving ①, ②

$x_1 = -7.686 \qquad x_2 = -8.101$

b) gradient

* Connection to output ①
(w_{01}, w_{11}, w_{12})

$\delta_1 = e_1 \cdot \frac{1}{2}(1 - s_1^2) = 0.588$

$\frac{\partial E}{\partial w_{01}} = -\delta_1 = -0.588$

$\frac{\partial E}{\partial w_{11}} = -\delta_1 x_1 = 4.519$

$\frac{\partial E}{\partial w_{12}} = -\delta_1 x_2 = 4.763$

* Connection to output ②

(w_{02}, w_{12}, w_{22})

$$\delta_2 = e_2 \cdot \frac{1}{2}(1 - s_2^2) = 0.324$$

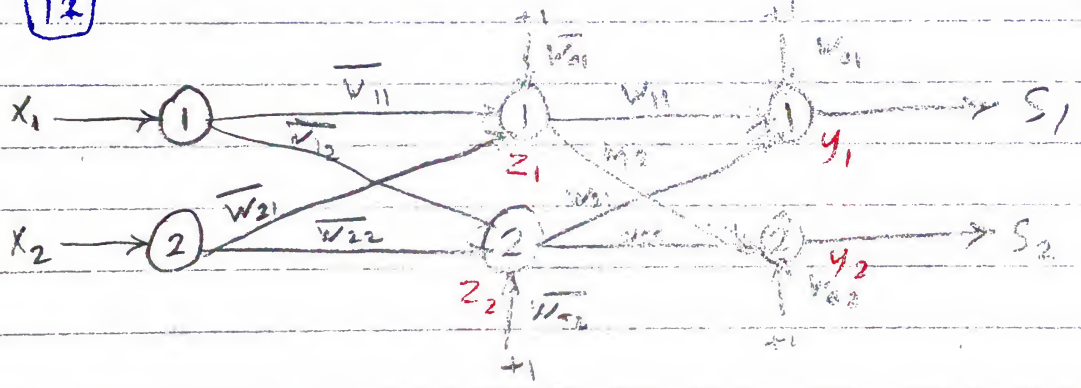
$$\frac{\partial E}{\partial w_{02}} = -\delta_2 = -0.324$$

$$\frac{\partial E}{\partial w_{12}} = -\delta_2 x_1 = -2.567$$

$$\frac{\partial E}{\partial w_{22}} = -\delta_2 x_2 = ~~-2.567~~ \; -2.706$$

12

Binary Sigmoid



Weights are given, $x_1 = -3.2$ $x_2 = 2.8$

$\delta_1 = 0.18$ $\delta_2 = 0.11$

a) find e_1, e_2

b) find gradient

a) ~~Fe~~ ~~o/p~~ ⓐ

For hidden ① → z_1

$z_1 = -1.164$

$F(z_1) = 0.238$

* For hidden ② z_2

$z_2 = 0.98$

$f(z_2) = 0.727$

* For output ① y_1

$y_1 = 1.103$

$f(y_1) = s_1 = 0.751$

* For output ② y_2

$y_2 = 1.003$

$f(y_2) = s_2 = -0.622$

$e_1 = d_1 - s_1 = -0.571$

$e_2 = d_2 - s_2 = -0.622$

---

b) gradient

* For output neuron ① (w_{01}, w_{11}, w_{21})

$\delta_1 = e_1 s_1'(1-s_1) = -0.107$

$\frac{\partial E}{\partial w_{01}} = -\delta_1 = 0.107$

$\frac{\partial E}{\partial w_{11}} = -\delta_1 f(x_1) = 0.25$

* For output ② (w_{02}, w_{12}, w_{22})

$\delta_2 = e_2 s_2 (1-s_2) = -0.122$

$\frac{\partial E}{\partial w_{02}} = -\delta_2 = 0.122$

$\frac{\partial E}{\partial w_{12}} = -\delta_2 f(x_2) = 0.29$

$$\frac{\partial E}{\partial w_{21}} = -\delta_1 f(z_2) = 0.078$$

$$\frac{\partial E}{\partial w_{22}} = -\delta_2 f(z_2) = 0.089$$

* for hidden ①

$$\frac{\partial E}{\partial w_{01}} = \left[\frac{\partial E}{\partial s_1} \cdot \frac{\partial s_1}{\partial y_1} \cdot \frac{\partial y_1}{\partial f(z_1)} \cdot \frac{\partial f(z_1)}{\partial z_1} \cdot \frac{\partial z_1}{\partial w_{01}}\right]$$

$$+ \left[\frac{\partial E}{\partial s_2} \cdot \frac{\partial s_2}{\partial y_2} \cdot \frac{\partial y_2}{\partial f(z_1)} \cdot \frac{\partial f(z_1)}{\partial z_1} \cdot \frac{\partial z_1}{\partial w_{01}}\right]$$

$$= \left[\underbrace{\frac{\partial E}{\partial s_1} \cdot \frac{\partial s_1}{\partial y_1}}_{-\delta_1} \cdot \underbrace{\frac{\partial y_1}{\partial f(z_1)}}_{w_{11}} + \underbrace{\frac{\partial E}{\partial s_2} \cdot \frac{\partial s_2}{\partial y_2}}_{-\delta_2} \cdot \underbrace{\frac{\partial y_2}{\partial f(z_1)}}_{w_{12}}\right] \frac{\partial f(z_1)}{\partial z_1} \cdot \underbrace{\frac{\partial z_1}{\partial w_{01}}}_{x_0}$$

$$= -\underbrace{\left[\delta_1 w_{11} + \delta_2 w_{12}\right] f(z_1)(1-f(z_1))}_{\delta_1} x_0$$

$$\frac{\partial E}{\partial w_{01}} = -\delta_1 x_0 = 0.029 \qquad \delta_1 = -0.029$$

$$\frac{\partial E}{\partial w_{11}} = -\delta_1 x_1 = -0.093$$

$$\frac{\partial E}{\partial w_{21}} = -\delta_1 x_2 = 0.081$$

* for hidden ②

$$\overline{\delta_2} = [\delta_1 w_{21} + \delta_2 w_{22}] F(z_2)(1 - F(z_2))$$

$$\overline{\delta_2} = -0.027$$

$$\frac{\partial E}{\partial \overline{w_{02}}} = -\overline{\delta_2} = 0.027$$

$$\frac{\partial E}{\partial \overline{w_{12}}} = -\overline{\delta_2} x_1 = -0.086$$

$$\frac{\partial E}{\partial \overline{w_{22}}} = -\overline{\delta_2} x_2 = 0.076$$